

Bounds on fast decodability of space-time block codes, skew-Hermitian matrices, and Azumaya algebras

Grégory Berhuy, Nadya Markin, and B. A. Sethuraman

Abstract

We study fast lattice decodability of space-time block codes for n transmit and receive antennas, written very generally as a linear combination $\sum_{i=1}^{2l} s_i A_i$, where the s_i are real information symbols and the A_i are $n \times n$ \mathbb{R} -linearly independent complex valued matrices. We show that the mutual orthogonality condition $A_i A_j^* + A_j A_i^* = 0$ for distinct basis matrices is not only sufficient but also necessary for fast decodability. We build on this to show that for full-rate ($l = n^2$) transmission, the decoding complexity can be no better than $|S|^{n^2+1}$, where $|S|$ is the size of the effective real signal constellation. We also show that for full-rate transmission, g -group decodability, as defined in [1], is impossible for any $g \geq 2$. We then use the theory of Azumaya algebras to derive bounds on the maximum number of groups into which the basis matrices can be partitioned so that the matrices in different groups are mutually orthogonal—a key measure of fast decodability. We show that in general, this maximum number is of the order of only the 2-adic value of n . In the case where the matrices A_i arise from a division algebra, which is most desirable for diversity, we show that the maximum number of groups is only 4. As a result, the decoding complexity for this case is no better than $|S|^{\lceil l/2 \rceil}$ for any rate l .

Index Terms

Fast Decodability, Full Diversity, Full Rate, Space-Time Code, Division Algebra, Azumaya Algebra.

I. INTRODUCTION

Space-time block codes for multiple input multiple output communications with n transmit and receive antennas and delay n and where the channel is known to the receiver consist of $n \times n$ matrices $X = X(x_1, \dots, x_l)$, $l \leq n^2$, where the symbols x_i arise from a finite subset S of the nonzero complex numbers. The matrices are generally assumed to be linear in the x_i , so splitting each x_i into its real and imaginary parts, we may write $X = \sum_{i=1}^{2l} s_i A_i$, where the s_i are real valued drawn from the effective real signal constellation S , and the A_i are fixed \mathbb{R} -linearly independent complex valued matrices. The transmission process may then be modeled as one where points from a $2l$ -dimensional lattice in \mathbb{R}^{2n^2} are transmitted (with the lattice changing every time the channel parameters change), and the decoding modeled as a closest lattice-point search.

Since closest lattice-point searches are notoriously difficult in general (although approximate decoding methods like sphere decoding [2] exist, which, by restricting the search points to a small region around the received point, speed up the process in small dimensions), much attention has been paid lately on selecting the matrices A_i above so that the resulting lattice breaks off as nearly as possible into an orthogonal direct sum of smaller dimensional lattices generated by some subsets of the canonical basis vectors, *no matter what the channel parameters* (see Remark 3 ahead for the interpretation of the previous work in terms

Grégory Berhuy is with Université Joseph Fourier, Institut Fourier, 100 rue des maths, BP 74, F-38402 Saint Martin d'Hères Cedex, France. E-mail: Gregory.Berhuy@ujf-grenoble.fr

Nadya Markin is with Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. E-mail: NMarkin@ntu.edu

B.A. Sethuraman is with Department of Mathematics, California State University Northridge, Northridge, CA 91330, USA. E-mail: al.sethuraman@csun.edu

of orthogonal sublattices). This then reduces the complexity of decoding from the worst case complexity $|S|^{2l}$ which arises from a brute-force checking of all $2l$ -tuples from S , to the order of $|S|^{l'}$ for some $l' < 2l$, where l' depends on the dimensions of the orthogonal summands. Some examples of recent work on fast decoding include [3], [4], [1], [6], [7], [8], [9], [10], [11]. Many codes have been shown to have reduced decoding complexity; for instance, it is known that the Silver code has a decoding complexity that is no higher than $|S|^5$ (instead of the possible $|S|^8$) [1, Example 5], considered in Example 2 ahead.

By *decoding complexity* we will mean throughout the complexity of the worst case decoding process whereby, upon possibly conditioning some variables, a brute-force check of the decoding metric is performed for all tuples from the remaining variables, possibly in parallel if the lattice has orthogonal direct summands. This is to be contrasted with other decoding processes that may exist that avoid brute force checking of the metric for all tuples, such as the GDL decoder described in [12].

In this paper, we analyze the conditions on the basis matrices A_i needed for reduced decoding complexity of space-time block codes arising from the phenomenon described above: the presence of orthogonal direct sums of smaller dimensional lattices generated by some subsets of the basis vectors of the transmitted lattice, no matter what the channel parameters. We show that the condition $A_i A_j^* + A_j A_i^* = 0$ for various distinct basis matrices A_i and A_j , previously considered in the literature primarily as a sufficient condition ([1] or [6] for instance, see also [4]), is actually a *necessary* condition (although, this result had indeed been proven before [5] using different techniques than ours, a fact we were unaware of: see Remark 1 ahead as well). We analyze this condition further, using just some elementary facts about skew-Hermitian and Hermitian matrices, and show that for a full-rate code (i.e., where $l = n^2$), the decoding complexity cannot be improved below $|S|^{n^2+1}$. We also show that for a full-rate code, the transmitted lattice cannot be decomposed entirely as an orthogonal direct sum of smaller dimensional lattices generated by the basis vectors (a condition referred to as *g-group decodability* by previous authors, for instance [1].)

We then drop the assumption of full rate and turn to the maximum number of orthogonal sublattices generated by basis vectors that is possible in the transmitted lattice; the dimension of the various sublattices then controls the fast-decodability. We use the theory of Azumaya algebras to show that the number of such summands is bounded above by $2v_2(n) + 4$ in general (where $v_2(n)$ is the 2-adic value of n , i.e., the highest power of 2 in the prime factorization of n). In the process, we generalize the classical Radon-Hurwitz-Eckmann bound [13] on the number of unitary matrices of square -1 that skew commute. Our method allows us to consider not just the general case but the special cases where the matrices A_i arise from embeddings of matrices over division algebras, where the bound on the number of summands becomes even smaller. In the case where the A_i come from the embedding of a division algebra, which is of most interest since codes from division algebras satisfy the full diversity criterion, we show that the maximum number of possible summands is very low: just 4 in fact. This then shows that the decoding complexity of a code arising from a division algebra cannot be made better than $|S|^{\lceil l/2 \rceil}$.

The paper is organized as follows: After some preliminary background on vectorizations of matrices and on Hermitian and skew-Hermitian matrices in Section II, we describe the system model and maximum likelihood decoding in Section III. We then discuss fast decodability in Section IV and derive the equivalence of fast decodability to the mutual orthogonality of subsets of the basis matrices. In Section V we analyze the mutual orthogonality condition using properties of skew-Hermitian and Hermitian matrices, and derive our lower bounds on the decoding complexity of full-rate codes. In Section VI, we use the theory of Azumaya Algebras to derive the bound on the number of orthogonal sublattices generated by basis vectors. Necessary background from commutative algebra and Azumaya algebras is collected in the appendices.

Acknowledgements: N. Markin was supported by the Singapore National Research Foundation under Research Grant NRF-RF2009-07. B.A. Sethuraman was supported by a U.S. National Science Foundation grant CCF-1318260. G. Berhuy and B.A. Sethuraman wish to thank Prof. Frederique Oggier and Nanyang Technological University, Singapore, for hosting their visit during which the ideas for this paper germinated. Portions of this paper were presented at the ISIT 2014 conference [14].

II. PRELIMINARIES

For any vector $\mathbf{v} \in \mathbb{C}^n$, we let

$$\text{Vec}_{\mathbb{R}}(\mathbf{v}) = (\text{Re}(v_1), \text{Im}(v_1), \dots, \text{Re}(v_n), \text{Im}(v_n))^t$$

be the vector in \mathbb{R}^{2n} whose $2i-1^{\text{th}}$ coordinate is the real part of v_i and whose $2i$ -th coordinate is the imaginary part of v_i . For any matrix $A \in M_n(\mathbb{C})$, we will write $\text{Vec}_{\mathbb{C}}(A)$ for the vector in \mathbb{C}^{n^2} obtained by stacking the entries of A in some fixed order (e.g. column 1 then column 2, etc.). To simplify notation, for a matrix A in $M_n(\mathbb{C})$, we will directly write $\text{Vec}_{\mathbb{R}}(A)$ for the vector $\text{Vec}_{\mathbb{R}}(\text{Vec}_{\mathbb{C}}(A))$ in \mathbb{R}^{2n^2} .

For two vectors \mathbf{v} and \mathbf{w} in \mathbb{C}^n , we write $\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{C}}$ for the usual Hermitian product in \mathbb{C}^n , namely, $\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{C}} = \mathbf{v} \cdot \mathbf{w}^* = \mathbf{v} \cdot \overline{\mathbf{w}}^t$ (where the superscript t stands for transpose). For two vectors \mathbf{v} and \mathbf{w} in \mathbb{R}^n , $\mathbf{v} \cdot \mathbf{w}$ will denote the dot product of the two vectors. For any matrix $A \in M_n(\mathbb{C})$, we will write A^* for the conjugate transpose of A , i.e., $A^* = \overline{A}^t$. Also, we will write Tr for the trace of a matrix, Re for the real part of a complex number.

The following are elementary:

Lemma 1. For two matrices A and B in $M_n(\mathbb{C})$, $\langle \text{Vec}_{\mathbb{C}}(A), \text{Vec}_{\mathbb{C}}(B) \rangle_{\mathbb{C}} = \text{Tr}(AB^*)$.

Lemma 2. For two vectors \mathbf{v} and \mathbf{w} in \mathbb{C}^n , $\text{Vec}_{\mathbb{R}}(\mathbf{v}) \cdot \text{Vec}_{\mathbb{R}}(\mathbf{w}) = \text{Re}(\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{C}})$.

We immediately get the following corollary:

Corollary 3. For two matrices A and B in $M_n(\mathbb{C})$, we have $\text{Vec}_{\mathbb{R}}(A) \cdot \text{Vec}_{\mathbb{R}}(B) = \text{Re}(\text{Tr}(AB^*))$. In particular, for matrices A and B , $\text{Vec}_{\mathbb{R}}(A)$ and $\text{Vec}_{\mathbb{R}}(B)$ are orthogonal in \mathbb{R}^{2n^2} if and only if $\text{Re}(\text{Tr}(AB^*)) = 0$.

We recall that a matrix $A \in M_n(\mathbb{C})$ is Hermitian if $A^* = A$, and skew-Hermitian if $A^* = -A$. The matrix $\imath I_n$ (where \imath is a square root of -1 and I_n is the identity $n \times n$ matrix) is skew-Hermitian. The set H_n of all Hermitian matrices and the set SH_n of all skew-Hermitian matrices in $M_n(\mathbb{C})$ each forms a vector space over \mathbb{R} , each of dimension n^2 . Moreover, for any Hermitian matrix A , $\imath A$ is skew-Hermitian, and for every skew-Hermitian matrix B , $\imath B$ is Hermitian. Every matrix can be written uniquely as a sum of a Hermitian and a skew-Hermitian matrix, i.e., $M_n(\mathbb{C}) \cong H_n \oplus SH_n$ as \mathbb{R} -vector spaces. We will need to use these facts in the paper.

III. SYSTEM MODEL AND MAXIMUM LIKELIHOOD DECODING

We consider transmission over a quasi-static Rayleigh fading channel with perfect channel state information at the receiver. We assume that the number of receive antennas and the number of transmit antennas are the same, namely n , and we assume the block length, i.e., the number of times we transmit through the channel before processing, is also n . The codewords are $n \times n$ complex valued matrices $X = X(x_1, \dots, x_l)$, $l \leq n^2$, where the symbols x_i arise from a finite subset of the nonzero complex numbers. The matrices X are assumed to be linear in the x_i , so splitting each x_i into its real and

imaginary parts, we may write $X = \sum_{i=1}^{2l} s_i A_i$, where the s_i are real symbols arising from the effective real alphabet S , and the A_i are fixed \mathbb{R} -linearly independent complex valued matrices. We will assume throughout the paper that the A_i are invertible, which is not a significant constraint, since invertible matrices form a dense subset of $n \times n$ complex matrices; besides, when the space-time code is fully diverse (which is the desirable situation), the matrices A_i are necessarily constrained to be invertible.

The received codeword is given by

$$Y = HX + N \tag{1}$$

where $H \in M_n(\mathbb{C})$ is the channel matrix and $N \in M_n(\mathbb{C})$ is the noise matrix. It is assumed that the entries of H are i.i.d. circularly symmetric complex Gaussian with zero mean and variance 1, and the entries of N are i.i.d. complex Gaussian with zero mean and variance N_0 .

The statistics of N shows that Maximum-likelihood (ML) decoding amounts to finding the information symbols s_1, \dots, s_{2l} that result in a codeword $X = \sum_{i=1}^{2l} s_i A_i$ which minimizes the squared Frobenius norm

$$\|Y - HX\|_F^2. \quad (2)$$

The transmission process may be modeled as one where points from a $2l$ -dimensional lattice in \mathbb{R}^{2n^2} are transmitted, with the lattice changing every time the channel matrix H changes, and the decoding modeled as a closest lattice-point search in \mathbb{R}^{2n^2} . We do this as follows: We convert the matrices appearing in Equation 1 to vectors in complex space and then further split the complex entries into their real and imaginary parts:

$$\text{Vec}_{\mathbb{R}}(Y) = \sum_{i=1}^{2l} s_i \text{Vec}_{\mathbb{R}}(H A_i) + \text{Vec}_{\mathbb{R}}(N).$$

We define $T = T(H)$ to be the $2n^2 \times 2l$ matrix over \mathbb{R} whose i -th column is $\text{Vec}_{\mathbb{R}}(H A_i)$. Then we have

$$\sum_{i=1}^{2l} s_i \text{Vec}_{\mathbb{R}}(H A_i) = T(s_1, \dots, s_{2l})^t$$

Thus, $T = T(H)$ is the basis matrix for the $2l$ -dimensional lattice in \mathbb{R}^{2n^2} from which points are transmitted. Writing \mathbf{s} for the vector $(s_1, \dots, s_{2l})^t$, the decoding problem now becomes to find a maximum likelihood estimate for the symbols s_1, \dots, s_{2l} from the linear system of equations in \mathbb{R}^{2n^2}

$$\text{Vec}_{\mathbb{R}}(Y) = T \cdot \mathbf{s} + \text{Vec}_{\mathbb{R}}(N), \quad (3)$$

where the entries of $\text{Vec}_{\mathbb{R}}(N)$ are i.i.d. real Gaussian. In other words, the decoding problem is to find an information vector $\mathbf{s} = (s_1, \dots, s_{2l})^t$ which minimizes the Euclidean distance

$$|\text{Vec}_{\mathbb{R}}(Y) - T\mathbf{s}| \quad (4)$$

of vectors in \mathbb{R}^{2n^2} .

Note that the transmitted lattice matrix $T = T(H)$ in Equation 3 above depends on the channel matrix H .

IV. FAST LATTICE DECODABILITY

Several authors ([3], [1]) studied fast lattice decodability of space-time codes by considering a QR decomposition of the transmitted lattice matrix T in Equation 3 above (as in the sphere decoder), and rewriting Equation 3 as

$$Q^* \text{Vec}_{\mathbb{R}}(Y) = R \cdot \mathbf{s} + Q^* \text{Vec}_{\mathbb{R}}(N). \quad (5)$$

Since Q^* is unitary, the new noise vector $Q^* \text{Vec}_{\mathbb{R}}(N)$ is still i.i.d. real Gaussian, so the maximum likelihood estimate for \mathbf{s} is given by minimizing $|Q^* \text{Vec}_{\mathbb{R}}(Y) - R \cdot \mathbf{s}|$. Fast lattice decodability as defined in [3], [1] involves choosing the basis matrices A_i so that for all H , the matrix R (which depends on $T(H)$ and hence on H), has zeros in certain convenient places (see Equation (6) ahead in the statement of Theorem 2, for instance). These places are such that decoding can proceed, after fixing certain s_i if necessary, as parallel decoding of smaller sets of variables, enabling thereby a reduction in complexity. We will study this process in this section, and prove the main result that enables us in the remaining sections to analyze bounds on fast decodability: the equivalence of fast decodability to mutual orthogonality of subsets of the basis matrices A_i (Theorem 5).

Definition 1. We say that two complex matrices, A, B are mutually orthogonal if $AB^* + BA^* = 0$.

We chose this term because, as we show in Theorem 1 below, two basis matrices A_i and A_j satisfy the relation $A_i A_j^* + A_j A_i^* = 0$ if and only if the i -th and j -th columns of T are mutually orthogonal as vectors in \mathbb{R}^{2l} . (Although our proof is new, see Remark 1 ahead.) The following lemma shows that mutually orthogonal matrices are necessarily \mathbb{R} -linearly independent:

Lemma 4. *If A_1, \dots, A_n are pairwise mutually orthogonal invertible matrices in $M_n(\mathbb{C})$, then they are \mathbb{R} -linearly independent.*

Proof: Assume that $r_1 A_1 + \dots + r_n A_n = 0$. Multiplying this equation on the right by A_i^* , and multiplying the conjugate transpose form of this equation on the left by A_i , and then adding, we find $2r_i A_i A_i^* = 0$. Since the A_i are invertible, we find $r_i = 0$. ■

Theorem 1. *The i -th and j -th columns of $T = T(H)$ are orthogonal as vectors in \mathbb{R}^{2l} for all channel matrices H if and only if the basis matrices A_i satisfy $A_i A_j^* + A_j A_i^* = 0$.*

Proof: We have already noted (Corollary 3 applied to the definition of the matrix T) that the orthogonality of the i -th and j -th columns of T is equivalent to the condition $\text{Re}(\text{Tr}((H A_i)(H A_j)^*)) = 0$. Also, note that $\text{Tr}((H A_i)(H A_j)^*) = \text{Tr}(H A_i A_j^* H^*) = \text{Tr}((A_i A_j^*)(H^* H))$, where the second equality is because $\text{Tr}(XY) = \text{Tr}(YX)$ for two matrices X and Y .

Now assume that $A_i A_j^* + A_j A_i^* = 0$ for $i \neq j$. Then $A_i A_j^*$ is skew-Hermitian, while $H^* H$ is of course Hermitian. If M is skew-Hermitian and P is Hermitian, then note that $(MP)^* = P^* M^* = -PM$. Since for any matrix X we have $\text{Re}(\text{Tr}(X)) = \text{Re}(\text{Tr}(X^*))$, we find that for $X = MP$, $\text{Re}(\text{Tr}(MP)) = \text{Re}(\text{Tr}((MP)^*)) = \text{Re}(\text{Tr}(-PM)) = -\text{Re}(\text{Tr}(PM)) = -\text{Re}(\text{Tr}(MP))$. It follows that $\text{Re}(\text{Tr}(MP)) = 0$. In particular, for $M = A_i A_j^*$ and $P = H^* H$, we find $0 = \text{Re}(\text{Tr}(A_i A_j^*)(H^* H)) = \text{Re}(\text{Tr}(H A_i)(A_j^* H^*)) = \text{Re}(\text{Tr}(H A_i)(H A_j)^*)$.

Now assume that the trace condition holds. We write this as $\text{Re}(\text{Tr}((A_i A_j^*)(H^* H))) = 0$ for all matrices H . Write M for $A_i A_j^*$. We wish to show that M is skew-Hermitian. The matrix $E_{k,k}$ that has 1 in the (k, k) slot and zeros elsewhere satisfies $E_{k,k}^* E_{k,k} = E_{k,k}$. Choosing $H = E_{k,k}$, we find that the matrix $M H^* H = M E_{k,k}$ will have the k -th column of M in the k -th column, and zeros elsewhere. The trace condition now shows that the (k, k) element of M is purely imaginary. We next need to show that $m_{l,k} = -\overline{m_{k,l}}$ for $k \neq l$, where we have written $m_{i,j}$ for the (i, j) -th entry of M . Computing directly, we find the following relations hold (where $E_{i,j}$ has 1 in the (i, j) slot and zeros everywhere else):

$$\begin{aligned} E_{k,k} + E_{k,l} + E_{l,k} + E_{l,l} &= (E_{k,k} + E_{l,k}) \cdot (E_{k,k} + E_{k,l}) \\ E_{k,k} - \imath E_{k,l} + \imath E_{l,k} + E_{l,l} &= (E_{k,k} + \imath E_{l,k}) \cdot (E_{k,k} - \imath E_{k,l}) \end{aligned}$$

Thus, each of the matrices on the left sides of the two equations above can be written as $H^* H$ for suitable matrices H . Again computing directly, we find that $M \cdot (E_{k,k} + E_{k,l} + E_{l,k} + E_{l,l})$ has $m_{k,k} + m_{k,l}$ in the (k, k) slot and $m_{l,k} + m_{l,l}$ in the (l, l) slot, and zeros elsewhere in the diagonal. Hence, $\text{Re}(\text{Tr}(M \cdot (E_{k,k} + E_{k,l} + E_{l,k} + E_{l,l}))) = \text{Re}(m_{k,k} + m_{k,l} + m_{l,k} + m_{l,l})$. Since we have already seen that the diagonal elements of M are purely imaginary, we find $\text{Re}(m_{k,l} + m_{l,k}) = 0$. Similarly, we find $\text{Re}(\text{Tr}(M \cdot (E_{k,k} - \imath E_{k,l} + \imath E_{l,k} + E_{l,l}))) = \text{Re}(m_{k,k} + \imath m_{k,l} - \imath m_{l,k} + m_{l,l})$. Once again, because the diagonal elements of M are purely imaginary, we find $\text{Im}(m_{k,l} - m_{l,k}) = 0$. These two together show that $m_{l,k} = -\overline{m_{k,l}}$ for $k \neq l$. Together with the fact that the diagonal elements of M are purely imaginary, we find $M = A_i A_j^*$ is skew-Hermitian, as desired. ■

Remark 1. As mentioned in Section I, the sufficiency of the condition $A_i A_j^* + A_j A_i^* = 0$ for orthogonality of the columns of T and hence for fast decodability was already considered before ([6, Theorem 2], [4, Theorem 1]). What is new here is the necessity of the condition. It is the consequences of the necessity that enables us to analyze lower bounds on fast decodability in the sections ahead by studying the consequences

of the condition $A_i A_j^* + A_j A_i^* = 0$. We should remark, however, that we noticed after we proved our results, that the authors of the paper [4] also mention the necessity of this condition. However, they do not give a proof of the necessity in that paper. Tracking this further, we discovered that the authors of [5] have actually provided a proof of this result. Their proof is by an explicit computation. Indeed, they write down the entries of $T(H)$, blockwise, in terms of the matrices H and A_i , and compute $T(H)^* T(H)$. From the derived block structure of $T(H)^* T(H)$ they read off the necessity of the mutual orthogonality. This is of course very different from our approach.

The theorem above allows us to define fast-decodability of a code in terms of its generating matrices, independently of the channel matrix H .

Definition 2. [See e.g., [1, Definition 5]] We will say that the space-time block code defined by the matrices $X = \sum_{i=1}^{2l} s_i A_i$ admits fast (lattice) decodability if for $g \geq 2$ there exist disjoint subsets $\Gamma_1, \dots, \Gamma_g, \Gamma_{g+1}$, with Γ_{g+1} possibly empty, of cardinalities n_1, \dots, n_g, n_{g+1} respectively, whose union is $\{1, \dots, 2l\}$, such that for all $u \in \Gamma_i$ and $v \in \Gamma_j$ ($1 \leq i < j \leq g$), the generating matrices A_u, A_v are mutually orthogonal.

Remark 2. Given a code that admits fast (lattice) decodability, we can define a permutation

$$\pi : \{1, \dots, 2l\} \rightarrow \Gamma_1 \cup \dots \cup \Gamma_g \cup \Gamma_{g+1},$$

which sends the first n_1 elements $\{1, \dots, n_1\}$ to Γ_1 , the next n_2 elements $\{n_1 + 1, \dots, n_1 + n_2\}$ to Γ_2 and so on, where, as in Definition 2, $n_i = |\Gamma_i|$ for $i = 1, \dots, g + 1$. Given such permutation π , we write T_π (or $T_\pi(H)$ for emphasized dependence on H) for the matrix whose i -th column is the $\pi(i)$ -th column of $T(H)$, namely, $\text{Vec}_{\mathbb{R}}(H A_{\pi(i)})$. Similarly, given the vector $\mathbf{s} = (s_1, \dots, s_{2l})^t$, we write \mathbf{s}_π for the vector whose i -th component is the $\pi(i)$ -th component of \mathbf{s} .

We are now able to link Definition 2 of fast-decodability to that given in [1, Definition 4]. While the latter definition invokes the channel matrix H , the two definitions are actually equivalent, for we have the following result:

Theorem 2. The space-time block code $X = \sum_{i=1}^{2l} s_i A_i$ admits fast (lattice) decodability as per Definition 2 if and only if there exists a permutation π of the index set $\{1, \dots, 2l\}$, integers $g \geq 2$, $n_i \geq 1$ ($i = 1, \dots, g$), and $n_{g+1} \geq 0$, with $n_1 + \dots + n_{g+1} = 2l$, such that for all channel matrices H , the matrix R obtained by doing a QR decomposition on $T_\pi = T_\pi(H)$ by doing a Gram-Schmidt orthogonalization in the order first column, then second column, and so on, has the special block form below:

$$\begin{pmatrix} B_1 & & & N_1 \\ & B_2 & & N_2 \\ & & \ddots & N_3 \\ & & & B_g & N_g \\ & & & & N_{g+1} \end{pmatrix} \quad (6)$$

for some matrices B_1, \dots, B_g , and N_1, \dots, N_{g+1} . Here, all empty spaces are filled by zeros, the B_i are of size $n_i \times n_i$ and the N_i are of size $n_i \times n_{g+1}$.

Before we prove this, we remark in more detail why previous authors have been interested in the special form of R above: On applying the permutation π to Equation 3, we get $\text{Vec}_{\mathbb{R}}(Y) = T_\pi \cdot \mathbf{s}_\pi + \text{Vec}_{\mathbb{R}}(N)$, and then, as in the beginning of this section, premultiplying by Q^* we find $Q^* \text{Vec}_{\mathbb{R}}(Y) = R \cdot \mathbf{s}_\pi + Q^* \text{Vec}_{\mathbb{R}}(N)$. It is clear from the block structure of the matrix R that after fixing the values of the last n_{g+1} variables in \mathbf{s}_π , the remaining variables can be decoded in g parallel steps, the i -th step involving n_i variables. The decoding complexity for this system is then of the order of $|S|^{n_{g+1} + \max n_i}$, where $|S|$ is the size of the effective real constellation S . This is in contrast to the complexity of $|S|^{2l}$ if the matrix R has no special structure.

Proof: If X is fast decodable as per Definition 2, then as described in Remark 2, the subsets $\Gamma_1, \dots, \Gamma_g, \Gamma_{g+1}$ provide a permutation π of $\{1, \dots, 2l\}$, and integers $g \geq 2$, n_1, \dots, n_g, n_{g+1} with the properties described.

Definition 2 and Theorem 1 also tell us that every column of T_π indexed by elements of $\pi^{-1}(\Gamma_i)$ is orthogonal to every column indexed by the elements of $\pi^{-1}(\Gamma_j)$ ($1 \leq i < j \leq g$). It follows immediately that on applying a QR decomposition to T_π in the order first column, then second column, etc., that the R matrix, which results from the Gram-Schmidt orthogonalizations of the columns of T_π in this order, will have the property that the columns indexed by $\pi^{-1}(\Gamma_i)$ will be perpendicular to those indexed by $\pi^{-1}(\Gamma_j)$. This can be seen easily from how the Gram-Schmidt process works, but this can also be checked from the explicit form of the matrix R obtained from this Gram-Schmidt orthogonalization, described for instance in [3, Section III] or [6, Section VI].

As for the other direction, assume that there is a permutation π of $\{1, \dots, 2l\}$ and integers $g \geq 2$, $n_i \geq 1$ ($i = 1, \dots, g$), and $n_{g+1} \geq 0$, with $n_1 + \dots + n_{g+1} = 2l$, such that for all H , $T_\pi(H) = QR$, where Q is unitary and R has the form as in Equation (6) above. Define the sets Γ_i in terms of the integers n_i as in Remark 2, namely $\Gamma_1 = \pi(\{1, \dots, n_1\})$ is the image of the first n_1 elements $\{1, \dots, n_1\}$, Γ_2 is the image of the next n_2 elements, and so on. It is clear from the block form of R that for any $u \in \Gamma_i$ and $v \in \Gamma_j$ ($1 \leq i < j \leq 2l$), the $\pi^{-1}(u)$ -th and $\pi^{-1}(v)$ -th columns of R are orthogonal as vectors in \mathbb{R}^{2n^2} . Since Q is unitary, the same holds for the matrix $T_\pi(H)$. Equivalently, the u -th and v -th columns of T are orthogonal for all H . Thus, by Theorem 1, A_u and A_v are mutually orthogonal, so X is fast decodable as per Definition 2. ■

We summarize what we have shown in the next corollary:

Corollary 5. *The following are equivalent for disjoint subsets $\Gamma_i, \Gamma_j \subset \{1, \dots, 2l\}$:*

- for all $u \in \Gamma_i$ and $v \in \Gamma_j$

$$A_u A_v^* + A_v A_u^* = 0.$$

- for all $u \in \Gamma_i$ and $v \in \Gamma_j$, the u -th and v -th columns of $T = T(H)$ are orthogonal as real vectors for any H .
- there exists a permutation π on the index set $\{1, \dots, 2l\}$ so that such that the matrix R arising as in the statement of Theorem 2 has a zero block in the entries $(\pi^{-1}(\Gamma_i), \pi^{-1}(\Gamma_j))$ and $(\pi^{-1}(\Gamma_j), \pi^{-1}(\Gamma_i))$.

Corollary 6. *Definition 2 of fast decodability is equivalent to one given in [1, Definition 4].*

Remark 3. In the notation of Definition 2, let L be the lattice in \mathbb{R}^{2n^2} generated by the columns of $T = T(H)$, and let L_i ($i = 1, \dots, g$) be the sublattices generated by the basis vectors of L coming from the columns in Γ_i (of the permuted matrix T_π). Fast-decodability can clearly be rephrased as the presence of sublattices L_i ($g \geq 2$) generated by subsets of the basis vectors that are orthogonal to one another in \mathbb{R}^{2n^2} . Indeed, previous work on fast decodability can be described in this language: seeking large numbers of sublattices generated by basis vectors that are orthogonal to one another.

Definition 3. *We say that the fast decodable code $X = \sum_{i=1}^{2l} s_i A_i$ is g -group decodable if it is fast (lattice) decodable and if Γ_{g+1} in Definition 2 is empty, so the matrix R of Theorem 2 has a block-diagonal form.*

Remark 4. As in the proof of Theorem 2, the block-diagonal structure of R of a g -decodable code translates (via pre-multiplication by Q) to the partitioning of the columns of T into g groups, the columns from any one group being orthogonal to the columns in any other group. Since T is the transmitted lattice matrix, we see that g -group decodability of the code is equivalent to the decomposition of the transmitted lattice into an orthogonal sum of smaller dimensional lattices generated by the basis vectors, no matter what the channel matrix H .

V. BOUNDS ON DECODING COMPLEXITY FOR FULL-RATE CODES

In this section, we will analyze the mutual orthogonality condition $A_i A_j^* + A_j A_i^* = 0$ of Theorem 5 and show that for full-rate codes, the best possible decoding complexity is not better than $|S|^{n^2+1}$ where $|S|$ is the size of the effective real constellation, and that g -group decoding is in fact not possible for full-rate codes. But first, we formalize the notion of decoding complexity:

Definition 4. The decoding complexity of the fast decodable space time code $X = \sum_{i=1}^{2l} s_i A_i$ is defined to be $|S|^{n_{g+1} + \max_{1 \leq i \leq g} n_i}$, where $n_i = |\Gamma_i|$, the Γ_i as in Definition 2.

Before delving into the main results of this section, we find it convenient to first gather a few lemmas concerning mutually orthogonal matrices that will be useful both here and in later sections.

Lemma 7. If matrices A and B are mutually orthogonal, so are MA and MB for any matrix M . If M is invertible, then A and B are mutually orthogonal if and only if MA and MB are mutually orthogonal.

Proof: This is a simple computation. ■

Lemma 8. If A and B are mutually orthogonal and A is invertible, then $A^{-1}B$ is skew-Hermitian.

Proof: By Lemma 7 above, $A^{-1}A = I_n$ and $A^{-1}B$ are mutually orthogonal. Writing down the mutual orthogonality condition for these two matrices, we find that $A^{-1}B$ is skew-Hermitian. ■

Lemma 9. The g invertible matrices $A_1 = I_n, A_2, \dots, A_g \in \mathcal{A} \subseteq M_n(\mathbb{C})$ are mutually orthogonal if and only if A_i is skew-Hermitian for $i \geq 2$ and A_2, \dots, A_g pairwise anticommute.

Proof: Assume that $A_1 = I_n, A_2, \dots, A_g \in \mathcal{A} \subseteq M_n(\mathbb{C})$ are mutually orthogonal. Since I_n and A_i are mutually orthogonal for $i \geq 2$, we find that A_i is skew-Hermitian for $i \geq 2$. In particular, for $i, j \geq 2, i \neq j$, we may replace A_i^* by $-A_i$ and A_j^* by $-A_j$ in the orthogonality relation to obtain the anticommuting relation $A_i A_j + A_j A_i = 0$. Conversely, assume that A_i is skew-Hermitian for $i \geq 2$ and A_2, \dots, A_g pairwise anticommute. We clearly have $I_n A_i^* + A_i I_n = 0$ for $i \geq 2$. Using the skew-Hermitian relation to replace the second factor in each summand of $A_i A_j + A_j A_i$ by the negative of its conjugate transpose, we find that the A_i , for $i = 2, \dots, g$ are mutually orthogonal. ■

Our first result is the following:

Theorem 3. Assume that the code $X = \sum_{i=1}^{2l} s_i A_i$ admits fast decodability, and let $k = \min_{1 \leq i \leq g} n_i$, where $n_i = |\Gamma_i|$, the Γ_i as in Definition 2. Then $n_1 + \dots + n_g \leq n^2 + k$.

Remark 5. In fact, we'll see later that if $k \geq 2$, then the sum $n_1 + \dots + n_g \leq n^2 + k - 1$.

We immediately get a high lower bound on the decoding complexity for full-rate codes from this theorem:

Corollary 10. The decoding complexity of a full-rate code of $n \times n$ matrices is at least $|S|^{n^2}$.

Proof: Since a full-rate code has exactly $2n^2$ basis matrices, this theorem shows that the subset Γ_{g+1} in Remark 2 must be of size at least $n^2 - k$, where $k = \min_{1 \leq i \leq g} n_i$. Having conditioned the symbols corresponding to Γ_{g+1} , decoding the first g groups of symbols in parallel has a decoding complexity at least $|S|^k$, therefore the decoding complexity of the entire code must be at least

$$|S|^{n^2-k} \cdot |S|^k = |S|^{n^2}.$$
■

We will show later that the bound is actually higher: it is $|S|^{n^2+1}$.

Corollary 11. *A full-rate code cannot be g -group decodable for $g \geq 3$.*

Proof: For, if a code is g -group decodable, then, written in the notation of Theorem 3, we have $2n^2 = n_1 + \dots + n_g \leq n^2 + k$, by the theorem. So $n^2 \leq k$, the number of elements in the smallest block, implying there can be at most 2 blocks. ■

We will see later that 2-group decodability is also not possible for full-rate codes.

We now prove the theorem.

Proof of Theorem 3: Let us denote the basis matrices in the groups Γ_i ($i = 1, \dots, g$) by $A_{i,j}$, $j = 1, \dots, n_i$. Multiplying the matrices on the left by any one $A_{i,j}^{-1}$ (recall from the beginning of Section III that we assume that the basis matrices are invertible), we replace one of the matrices in the i -th block by the identity matrix I_n , and as for the modified matrices in the other blocks, they are now orthogonal to I_n by Lemma 7 above. By Lemma 8 above, the modified matrices $A_{i,j}^{-1}A_{k,l}$ in the remaining blocks are all skew-Hermitian as well. Since the remaining matrices $A_{i,j}^{-1}A_{k,l}$ are also \mathbb{R} -linearly independent by Lemma 4, and since the dimension of the space of skew-Hermitian $n \times n$ matrices over \mathbb{R} is n^2 (Section II), we find that for each i , $(n_1 + \dots + n_g) - n_i \leq n^2$. The result now follows immediately. ■

Our next few results will help us sharpen the bounds on decoding complexity we obtain from Theorem 3 (see Corollary 10).

Theorem 4. *There can be at most $n^2 - 1$ \mathbb{R} -linearly independent matrices in $M_n(\mathbb{C})$ that are both skew-Hermitian and mutually orthogonal.*

Proof: For, suppose to the contrary that A_1, \dots, A_{n^2} were \mathbb{R} -linearly independent, skew-Hermitian, and mutually orthogonal. The matrix ιI_n is skew-Hermitian. Suppose first that one of these A_i , say A_1 , is an \mathbb{R} -multiple of ιI_n . This is already a contradiction, since $A_1 A_2^*$ is skew-Hermitian by the mutual orthogonality condition, but $A_1 A_2^*$ is a real multiple of ιA_2^* and is therefore Hermitian. Now suppose that no A_i is an \mathbb{R} -multiple of ιI_n . The matrix ιI_n , being skew-Hermitian, can be written as a linear combination of these matrices A_i since they form a basis for the skew-Hermitian matrices, so $\iota I_n = \sum a_j A_j$ for real a_j . Now A_1 is not a real multiple of ιI_n by assumption. Consider $\iota I_n A_1^*$. This is Hermitian. On the other hand, $(\sum a_j A_j) A_1^* = a_1 A_1 A_1^* + (\sum_{j=2} a_j A_j) A_1^*$, where this second sum runs from $j = 2$ onwards. But for $j = 2$ onwards, $A_j A_1^*$ is skew-Hermitian by the mutual orthogonality condition, while both ιA_1^* and $a_1 A_1 A_1^*$ are Hermitian. For this to happen, $(\sum_{j=2} a_j A_j) A_1^*$, where the sum is over $j \geq 2$, must be zero, and ιA_1^* must equal $a_1 A_1 A_1^*$. On canceling A_1^* (recall our assumption that the basis matrices are invertible), we find that A_1 is a multiple of ιI_n , contradiction. ■

Example 1. In the 2×2 matrices $M_2(\mathbb{C})$ over the complex numbers \mathbb{C} , consider the three matrices $A_1 = \begin{pmatrix} \iota & 0 \\ 0 & -\iota \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and $A_3 = \begin{pmatrix} 0 & -\iota \\ -\iota & 0 \end{pmatrix}$. These three matrices are \mathbb{R} -linearly independent, skew-Hermitian, and pairwise mutually orthogonal matrices. Together with the identity matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, they form a \mathbb{C} -basis for $M_2(\mathbb{C})$, and as can be checked, no \mathbb{C} -linear combination of I , A_1 , A_2 , and A_3 is both skew-Hermitian and mutually orthogonal to A_1 , A_2 , and A_3 . Thus, the $2^2 - 1$ matrices A_1 , A_2 , and A_3 exemplify the contention of this theorem.

We get a quick corollary from this that we will sharpen considerably in the next section:

Corollary 12 (See Corollary 16 in Section VI). *For a code generated by invertible $n \times n$ matrices, the maximum number of groups g in notation of Definition 2 is n^2 .*

Proof: If the number of groups is more than n^2 , then we can find $n^2 + 1$ matrices that are \mathbb{R} -linearly independent and mutually orthogonal. Multiplying this set on the left by the inverse of one of them (as in the proof of Theorem 3 above), we find n^2 skew-Hermitian and mutually orthogonal \mathbb{R} -linearly independent matrices, a contradiction. ■

Lemma 13. *If any $g - 1$ of the groups $\Gamma_1, \dots, \Gamma_g$ from Definition 2 together have at least n^2 matrices in them, then they have exactly n^2 elements in them, while the remaining group can only have one matrix in it.*

Proof: Say the last $g - 1$ groups, for simplicity, together have at least n^2 matrices, and suppose that the first group has at least two elements, call them A and B . By multiplying throughout by A^{-1} , we can assume that the two elements are I and B . Note that after multiplying by A^{-1} , because of the mutual orthogonality condition, the matrices in the remaining groups all become skew-Hermitian (as in the proof of Theorem 3 above). Because there are at least n^2 skew-Hermitian (\mathbb{R} -linearly independent) matrices, we find that there must be exactly n^2 of them because the dimension of the skew-Hermitian matrices is n^2 . Call these n^2 matrices C_1, \dots, C_{n^2} . We must have $\imath I_n$ in the linear span of these C_i because $\imath I_n$ is also skew-Hermitian. Thus, $\imath I_n = \sum a_i C_i$. Now multiply on the right by B^* , where B is as above. Each of the products $C_i B^*$ is skew-Hermitian because of the mutual orthogonality condition that requires $C_i B^* + B C_i^* = 0$. Thus, $\imath B^*$ is also skew-Hermitian. It follows from this that B^* is Hermitian, i.e., B is Hermitian. But now, we consider $C_i B^*$ for any i . The mutual orthogonality condition says that this is skew-Hermitian, so it equals $-(B C_i^*)$, and since C_i^* is skew-Hermitian, this equals $B C_i$. On the other hand, we just saw that B is Hermitian, so $C_i B^* = C_i B$. Thus, B commutes with all C_i , i.e., with all skew-Hermitian matrices. But this means B commutes with all the Hermitian matrices as well, because every Hermitian matrix is of the form \imath times a skew-Hermitian matrix. Thus, B commutes with all matrices, and is Hermitian, so it must be a real scalar matrix. But this violates the fact that I_n and B were two linearly independent matrices in the first group. ■

Corollary 14. *If, as in the notation of Definition 2, $n_i \geq 2$ for any i , then the total number of matrices in the g groups is at most $n^2 + n_i - 1$. In particular, if $k = \min n_i \geq 2$, then the total number is at most $n^2 - 1 + k$.*

Proof: Since the i -th group has size $n_i \geq 2$, the remaining groups must have less than n^2 matrices in them, or else, the lemma above will be violated. It follows that there at most $n^2 + n_i - 1$ matrices in the g groups. ■

We are now ready to sharpen the results we got in Corollary 10.

Theorem 5. *The decoding complexity of a full-rate space time code $X = \sum_{i=1}^{2n^2} s_i A_i$ is not better than $|S|^{n^2+1}$, where $|S|$ is the size of the effective real constellation.*

Proof: Consider the basis matrices A_i : if there are at least two mutually orthogonal groups, then, by Definition 2, the code is fast decodable, and by Theorem 2 the R matrix that comes from $T = T(H)$ will have the form (6). Consider the integers n_i , notation as in Definition 2. If any $n_i \geq 2$, then by Corollary 14, the total number of matrices in the g groups is at most $n^2 + n_i - 1$. Thus, the matrix N_{g+1} in (6) will be of size at least $(n^2 - n_i + 1) \times (n^2 - n_i + 1)$. Exactly as in the proof of Corollary 10, we find that the decoding complexity must be at least $|S|^{n^2 - n_i + 1} \cdot |S|^{n_i} = |S|^{n^2 + 1}$. If on the other hand all $n_i = 1$, then we have g groups of size 1 each. By Corollary 12, $g \leq n^2$, so N_{g+1} is at least of size $n^2 \times n^2$. Thus, there are at least n^2 variables corresponding to N_{g+1} that need to be conditioned, and then, the g blocks are decoded in parallel, with complexity $|S|$ each. Thus the decoding complexity is at least $|S|^{n^2} \cdot |S| = |S|^{n^2 + 1}$. ■

Example 2. Silver Code: This 2×2 code for four complex signal elements s_1, s_2, s_3, s_4 is given by $X(s_1, s_2) + T X(z_1, z_2)$, where for any a and b , $X(a, b) = \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix}$, and $T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The signal elements s_3 and s_4 are related to z_1 and z_2 by $(z_1, z_2)^T = M(s_3, s_4)^T$, where $M = \frac{1}{\sqrt{7}} \begin{pmatrix} 1 + \imath & -1 + 2\imath \\ 1 + 2\imath & 1 - \imath \end{pmatrix}$.

This code has a decoding complexity of at most $|S|^5$ (see [1] for instance). This example thus shows that our bound $n^2 + 1$ is strict. Moreover, Theorem 5 shows that the Silver code cannot have a lower lattice

decoding complexity than the known $|S|^5$.

Theorem 6. *It is not possible to arrange for the full-rate space-time code $X = \sum_{i=1}^{2n^2} s_i A_i$ to have g -group decodability for any g .*

Proof: We have already seen in Corollary 11 that g -group decodability is not possible for $g \geq 3$. For $g = 2$, note that one of two groups must have at least n^2 matrices. It follows from Lemma 13 that this group must have exactly n^2 elements and the other group must have only one element. Since $n \geq 2$ in the space-time block code paradigm, $1 + n^2 < 2n^2$, and 2-group decodability is hence impossible. ■

Remark 6. In a different language (see Remark 4), Theorem 6 says that the transmitted lattice of a full-rate space-time code does not split off as an orthogonal sum of smaller dimensional lattices generated by the canonical basis vectors.

VI. AZUMAYA ALGEBRAS AND BOUNDS ON THE NUMBER OF GROUPS

In this section, we will delve into the arithmetic of central-simple algebras, using machinery from commutative ring theory and Azumaya algebras, to determine significantly small upper bounds on the number of orthogonal sublattices generated by the basis vectors of the transmitted lattice $T = T(H)$, or what is the same, the number of blocks g of the R matrix in Equation (6). We had already derived an upper bound of n^2 for full-rate codes in Corollary 12, but as we will see, this bound is too high. In fact, the bound behaves more like $\log_2(n)$ (see Theorem 8 for a precise statement). The bound we derive in this section will be independent of the code rate (l). Since the matrices in distinct groups are pairwise mutually orthogonal, we will derive our bound by answering the following question: How many \mathbb{R} -linearly independent pairwise mutually orthogonal matrices can we find in $M_n(\mathbb{C})$? In fact, we will actually answer a broader question: Let $k \subset \mathbb{C}$ be a number field, let \mathcal{A} be a central simple k -subalgebra of $M_n(\mathbb{C})$. How many \mathbb{R} -linearly independent pairwise mutually orthogonal matrices can we find in the subalgebra $\mathcal{A} \subseteq M_n(\mathbb{C})$? (Of course, by Lemma 4, we may drop the requirement that the matrices be \mathbb{R} -linearly independent.)

As in the earlier sections, we will assume that our pairwise orthogonal matrices are all invertible. Note that if a matrix $A \in \mathcal{A} \subseteq M_n(\mathbb{C})$ is invertible as an element of $M_n(\mathbb{C})$, its inverse must actually lie in \mathcal{A} . This is because A^{-1} can be obtained from the minimal polynomial of A over k as follows: if the minimal polynomial is $A^t + k_{t-1}A^{t-1} + \cdots + k_1A + k_0$, then $k_0 \neq 0$ because A is invertible as a matrix, so the inverse of A can be written by factoring out A as $(-1/k_0)(A^{t-1} + k_{t-1}A^{t-2} + \cdots + k_1)$. The inverse of A hence lives in the subalgebra $k[A] \subseteq \mathcal{A}$.

All the k -algebras we consider will be implicitly assumed to be finite-dimensional over k . Various background facts about commutative rings and Azumaya algebras are collected in Appendices A and B respectively. We will assume basic knowledge of central simple algebras (see [15] for instance).

Lemmas 7, 8, and 9 show us that the existence of (invertible) mutually orthogonal matrices A_i , $i = 1, \dots, m$ is equivalent (upon replacing the A_i by say $A_1^{-1}A_i$) to the existence of matrices $C_i = A_1^{-1}A_i$, $i = 2, \dots, m$ which are skew-Hermitian and anticommute pairwise.

So, focusing on the necessary anticommuting condition above, we study the following question. (In the sequel, \mathcal{A}^\times will refer to the invertible elements of \mathcal{A} .)

Question. Let k be a number field, and let \mathcal{A} be a central simple k -algebra. How many elements $u_1, \dots, u_r \in \mathcal{A}^\times$ which pairwise anticommute can we find?

We now investigate this question.

Once and for all, we fix a central simple k -algebra \mathcal{A} , and we assume to have elements $u_1, \dots, u_r \in \mathcal{A}^\times$ such that $u_i u_j + u_j u_i = 0$ for all $i \neq j$, for some $r \geq 2$. For the moment, we only assume that k is any field of characteristic different from 2.

Notice that u_i and u_j^2 commute for all i, j . Indeed, this is clear if $i = j$, and if $i \neq j$, we have $u_i u_j^2 = -u_j u_i u_j = u_j^2 u_i$.

This implies that u_i^2, u_j^2 commute for all i, j . Consequently, the k -algebra

$$R = k[u_1^2, u_1^{-2}, \dots, u_r^2, u_r^{-2}]$$

is a commutative k -subalgebra of \mathcal{A} . (Of course, as remarked in the second paragraph of this section, the k -algebra generated by u_i^2 will already contain u_i^{-2} , but we choose to include the u_i^{-2} in the generators of R to emphasize that the u_i^2 are units in R , a fact we will need below.)

Notice also that for any u_{i_1}, \dots, u_{i_k} , we have $(u_{i_1} \cdots u_{i_k})^2 = \pm u_{i_1}^2 \cdots u_{i_k}^2 \in R^\times$.

We recall the definition of the algebra $(a, b)_R$ from Part 4 of Examples 6 in Appendix B: given a commutative ring R and a, b in R^\times , $(a, b)_R$ is the R -algebra generated by two elements e and f subject to the relations $e^2 = a$, $f^2 = b$, and $fe = -ef$. It has the matrix realization described in Appendix B.

Lemma 15. *Let $r = 2s$ or $2s + 1$. Keeping notation above, \mathcal{A} contains a subring isomorphic to*

$$(a_1, b_1)_R \otimes_R \cdots \otimes_R (a_s, b_s)_R,$$

for some $a_p, b_p \in R^\times$.

Proof: If I is any subset of $\{1, \dots, n\}$, set $u_I = \prod_{i \in I} u_i$. It is then easy to check that for all I, J , we

$$\text{have } u_I u_J = (-1)^{|I| \cdot |J| - |I \cap J|} u_J u_I.$$

For $p = 1, \dots, s$, set

$$I_p = \{1, \dots, 2p - 1\}, \quad J_p = \{1, \dots, 2p - 2, 2p\}.$$

We then have $|I_p| = |J_p| = 2p - 1$, $|I_p \cap J_p| = 2p - 2$, and for all $1 \leq p < q \leq s$, we have $|I_p \cap I_q| = |J_p \cap J_q| = |I_p \cap J_q| = |I_q \cap J_p| = 2p - 1$.

Now set

$$\alpha_p = u_{I_p}, \quad \beta_p = u_{J_p}.$$

Notice that $a_p = \alpha_p^2, b_p = \beta_p^2 \in R^\times$. Moreover, for all $p = 1, \dots, s$, we have $\alpha_p \beta_p = u_{I_p} u_{J_p} = (-1)^{(2p-1)^2 - (2p-2)} u_{J_p} u_{I_p} = -u_{J_p} u_{I_p} = -\beta_p \alpha_p$. Thus, for all $p = 1, \dots, s$, we have an R -algebra morphism $\varphi_p : (a_p, b_p)_R \rightarrow \mathcal{A}$, which maps the generators e_p and f_p onto α_p and β_p respectively.

Now for all $1 \leq p < q \leq s$, we have $\alpha_p \alpha_q = u_{I_p} u_{I_q} = (-1)^{(2p-1)(2q-1) - (2p-1)} u_{J_p} u_{I_p} = \alpha_q \alpha_p$. Similarly, we have $\beta_p \beta_q = \beta_q \beta_p$. We also have $\alpha_p \beta_q = u_{I_p} u_{J_q} = (-1)^{(2p-1)(2q-1) - (2p-1)} u_{J_q} u_{I_p} = \beta_q \alpha_p$. Similarly, we have $\alpha_q \beta_p = \beta_p \alpha_q$.

It follows that $\varphi_1, \dots, \varphi_s$ have pairwise commuting images. Thus, they induce an R -algebra morphism

$$(a_1, b_1)_R \otimes_R \cdots \otimes_R (a_s, b_s)_R \rightarrow \mathcal{A}.$$

By Lemma 20 and Remark 9, this morphism is injective. ■

We may now give a full answer to the previous question.

Theorem 7. *Let k be a number field, and let \mathcal{A} be a central simple k -algebra. Let u_1, \dots, u_r ($r \geq 2$) be invertible elements in \mathcal{A} which pairwise anticommute. Then we have*

$$r \leq 2\nu_2 \left(\frac{\deg(\mathcal{A})}{\text{ind}(\mathcal{A})} \right) + 2 \text{ if } r \text{ is even}$$

and

$$r \leq 2\nu_2 \left(\frac{\deg(\mathcal{A})}{\text{ind}(\mathcal{A})} \right) + 3 \text{ if } r \text{ is odd,}$$

where ν_2 denotes the 2-adic value of an integer, i.e., the highest power of 2 that divides that integer.

In particular, if \mathcal{A} is a central division k -algebra, then $r = 2, 3$.

Remark 7. See Appendix C for how this result above compares with the classical Hurwitz-Radon-Eckmann bound on anticommuting matrices.

Proof: We may assume $r > 2$ if r is even, and $r > 3$ if r is odd, since otherwise this is trivial. Write $r = 2s$ or $r = 2s + 1$, so $s \geq 2$. By the previous lemma, \mathcal{A} contains an R -algebra isomorphic to

$$(a_1, b_1)_R \otimes_R \cdots \otimes_R (a_s, b_s)_R,$$

for some $a_p, b_p \in R^\times$. By Proposition 22 in Appendix B applied $s - 1$ times (note that $s - 1 \geq 1$ by assumption), this R -algebra is isomorphic to

$$M_{2^{s-1}}(R) \otimes_R (c, d)_R \cong_R (M_{2^{s-1}}(k) \otimes_k R) \otimes_R (c, d)_R$$

for some $c, d \in R^\times$. Hence \mathcal{A} contains a k -subalgebra isomorphic to $M_{2^{s-1}}(k)$. The centralizer theorem then implies that

$$\mathcal{A} \cong_k M_{2^{s-1}}(k) \otimes_k \mathcal{A}',$$

for some central simple k -algebra \mathcal{A}' , which is Brauer-equivalent to \mathcal{A} by definition. Therefore, we may write

$$\mathcal{A} \cong_k M_\ell(D), \quad \mathcal{A}' \cong_k M_t(D),$$

where D is a central division k -algebra. Thus, we get

$$M_\ell(D) \cong_k M_{2^{s-1}t}(D),$$

and then $2^{s-1}t = \ell = \frac{\deg(\mathcal{A})}{\text{ind}(\mathcal{A})}$. The desired result follows easily. \blacksquare

Remark 8. If \mathcal{A} is a central simple k -algebra of odd degree, then \mathcal{A} does not contain pairwise anticommuting invertible elements.

Indeed, if u_1 and u_2 anticommute, then we have

$$\text{Nrd}_{\mathcal{A}}(u_1 u_2) = \text{Nrd}_{\mathcal{A}}(u_1) \text{Nrd}_{\mathcal{A}}(u_2) = \text{Nrd}_{\mathcal{A}}(-u_2 u_1) = -\text{Nrd}_{\mathcal{A}}(u_2) \text{Nrd}_{\mathcal{A}}(u_1),$$

where the last equality arises from the fact that $\text{Nrd}_{\mathcal{A}}(-1) = -1$ since \mathcal{A} has odd degree. Hence, $\text{Nrd}_{\mathcal{A}}(u_1) \text{Nrd}_{\mathcal{A}}(u_2) = 0$. But the reduced norm of an invertible element of \mathcal{A} is non-zero, hence a contradiction.

Hence the previous bounds are not always sharp. However they may be sharp in certain cases as the following example shows, which proves that these bounds are the best possible ones.

Example 3. Let $\ell \geq 0$ be an integer, let $Q = (a, b)_k$ be a division quaternion k -algebra, and let $\mathcal{A} = M_{2^\ell}(k) \otimes_k Q$.

In order to avoid mixing notation, we will denote exceptionally by \odot the Kronecker product of matrices. If $t \geq 0$ is an integer, we denote by $M^{\odot t}$ the Kronecker product of t copies of M , where $M^{\odot 0}$ is the identity matrix by convention.

Let

$$H_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad H_{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

For $p = 1, \dots, \ell$, set

$$U_{2p-1} = H_1^{\odot(p-1)} \odot H_1 H_{-1} \odot I_2^{\odot(\ell-p)} \quad \text{and} \quad U_{2p} = H_1^{\odot(p-1)} \odot H_{-1} \odot I_2^{\odot(\ell-p)}.$$

The properties of the Kronecker product and the fact that $H_1 H_{-1} = -H_{-1} H_1$, show that U_1, \dots, U_{2p} are invertible matrices of $M_{2^\ell}(k)$ which pairwise anticommute.

Now let e and f be the generators of Q . Then it is easy to check that the $2\ell + 3$ invertible elements

$$U_1 \otimes 1, \dots, U_{2\ell} \otimes 1, U_1 \cdots U_{2\ell} \otimes e, U_1 \cdots U_{2\ell} \otimes f, U_1 \cdots U_{2\ell} \otimes ef \in \mathcal{A}$$

pairwise anticommute.

Notice for later use that U_{2p-1} is symmetric and U_{2p} is skew-symmetric for $p = 1, \dots, \ell$. Notice also that $U_1 \cdots U_{2\ell}$ is symmetric, as a straightforward computation shows.

As a corollary, we also get an answer to our main problem.

Corollary 16. *Let k be a number field, let \mathcal{A} be a central simple k -subalgebra of $M_n(\mathbb{C})$. Assume that we have g non-zero matrices $A_1, \dots, A_g \in \mathcal{A}^\times$ ($g \geq 2$) such that*

$$A_i^* A_j + A_j^* A_i = 0 \text{ for all } i \neq j.$$

Then $g \leq 2\nu_2(\frac{\deg(\mathcal{A})}{\text{ind}(\mathcal{A})}) + 3$ if g is odd, and $g \leq 2\nu_2(\frac{\deg(\mathcal{A})}{\text{ind}(\mathcal{A})}) + 4$ if g is even.

In particular, if \mathcal{A} is a central division k -algebra, then $g \leq 4$.

Proof: By Lemma 9, the existence of g such matrices implies the existence of $g-1$ invertible elements of \mathcal{A} which pairwise anticommute. Now apply the previous theorem to conclude. ■

The next example shows that these bounds may be sharp.

Example 4. Let $k \subset \mathbb{R}$, and let $U_1, \dots, U_{2\ell} \in M_{2\ell}(k) \subset M_{2\ell}(\mathbb{R})$ be the matrices introduced in Example 3. Set $Q = (-1, -1)_k$, so that Q is a division k -algebra.

The multiplication matrices of e and f with respect to the $k(i)$ -basis $(1, e)$ of Q (viewed as a right $k(i)$ -vector space) are the skew-Hermitian matrix iH_1 and the hermitian matrix H_{-1} respectively. Notice that $iH_1 H_{-1}$ is skew-Hermitian. The results of Example 3 show that the matrices

$$U_1 \odot I_2, \dots, U_{2\ell} \odot I_2, U_1 \cdots U_{2\ell} \odot (iH_1), U_1 \cdots U_{2\ell} \odot H_{-1}, U_1 \cdots U_{2\ell} \odot (iH_1 H_{-1})$$

pairwise anticommute.

Each of these matrices are hermitian or skew-Hermitian. Multiplying by i the appropriate matrices yields a set of $2\ell + 3$ skew-Hermitian matrices which pairwise anticommute. More precisely, one may check that the matrices

$$U_{2p-1} \odot I_2, U_{2p} \odot (iI_2), p = 1, \dots, \ell,$$

$$U_1 \cdots U_{2\ell} \odot (iH_1), U_1 \cdots U_{2\ell} \odot (iH_{-1}), U_1 \cdots U_{2\ell} \odot (iH_1 H_{-1})$$

are skew-Hermitian matrices which pairwise anticommute. Adding the identity matrix then gives rise to a set of $2\ell + 4$ mutually orthogonal matrices.

It is worth rewording the result in Corollary 16 in the language of our space-time code. We have the following:

Theorem 8. *If the space-time code $X = \sum_{i=1}^{2l} s_i A_i$ is fast-decodable, then the number of groups g in (6) is at most $2\nu_2(n) + 4$. If we assume that the A_i are chosen from some k -central simple algebra $\mathcal{A} \subseteq M_n(\mathbb{C})$, where k is some number field, then, this upper bound drops to $g \leq 2\nu_2(\frac{\deg(\mathcal{A})}{\text{ind}(\mathcal{A})}) + 4$. In particular, if the A_i are chosen from a k -central division algebra, then $g \leq 4$.*

We get an immediate corollary:

Corollary 17. *The decoding complexity of a fast decodable space-time code $X = \sum_{i=1}^{2l} s_i A_i$ where the A_i are chosen from a division algebra is at least $|S|^{\lceil l/2 \rceil}$.*

Proof: At least one of the groups Γ_i ($i = 1, \dots, g$) in Definition 2 must be of size at least $\lceil 2l/4 \rceil$, as $g \leq 4$ when the A_i are chosen from a division algebra. Thus, the decoding complexity is at least $|S|^{n_{g+1} + \lceil 2l/4 \rceil} \geq |S|^{\lceil l/2 \rceil}$. ■

APPENDIX A COMMUTATIVE ALGEBRA

We collect here some useful results in commutative algebra. We start with the notion of an Artin ring.

Definition 5. A commutative ring R is an Artin ring if every descending chain of ideals $I_0 \supset I_1 \supset I_2 \supset \dots$ of R is stationary, i.e., there exists $n > 0$ such that $I_n = I_{n+1} = I_{n+2} = \dots$.

Example 5. If k is a field, any finite-dimensional commutative k -algebra R is an Artin ring. Indeed, any ideal is in particular a finite-dimensional k -subspace of R , so it cannot exist a strictly decreasing chain of ideals.

Theorem 9. [18, Ch.8, Thm 8.5] Any Artin ring is Noetherian, that is every ideal is finitely generated.

Corollary 18. Let R be a local Artin ring, with maximal ideal \mathfrak{m} . Then there exists $n \geq 1$ such that $\mathfrak{m}^n = 0$.

Proof: By assumption, the descending chain of ideals $\mathfrak{m} \supset \mathfrak{m}^2 \supset \dots \supset \mathfrak{m}^n \supset \dots$ is stationary, hence there exists $n \geq 1$ such that $\mathfrak{m}^{n+1} = \mathfrak{m} \cdot \mathfrak{m}^n = \mathfrak{m}^n$. Since R is Noetherian by the previous theorem, \mathfrak{m} is finitely generated, and since R is local with unique maximal ideal \mathfrak{m} , $\mathfrak{m}^n = 0$ by Nakayama's lemma. ■

We also have the following result.

Theorem 10. [18, Ch.8, Thm. 8.7] Any Artin ring is isomorphic to the direct product of finitely many Artin local rings. In particular, an Artin ring has finitely many maximal ideals.

We now define Hensel rings.

Definition 6. A commutative ring R is a Hensel ring if R is local, with maximal ideal \mathfrak{m} , and for any monic polynomial $f \in R[X]$ such that $\bar{f} = \bar{g}_0 \bar{h}_0 \in R/\mathfrak{m}[X]$ for some coprime monic polynomials $\bar{g}_0, \bar{h}_0 \in R/\mathfrak{m}[X]$, there exists coprime monic polynomials $g, h \in R[X]$ such that $f = gh$ and $\bar{g} = \bar{g}_0, \bar{h} = \bar{h}_0$.

The following result is well-known.

Proposition 19. Any local Artin ring is a Hensel ring.

Proof: Since the maximal ideal \mathfrak{m} of a local ring is nilpotent by Corollary 18, R is canonically isomorphic to its \mathfrak{m} -completion, that is R is complete. Since complete rings are Hensel rings by [19, Prop. 4.5], we are done. ■

APPENDIX B AZUMAYA ALGEBRAS

We collect here some notions on Azumaya algebras that are needed in the paper. The word ‘algebra’ implicitly means ‘associative algebra with unit’.

In this section, R is a commutative ring with unit. We first define Azumaya R -algebras. The reader willing to learn more about Azumaya algebras will refer to [20, III.5].

Definition 7. An Azumaya R -algebra is an R -algebra A , which is finitely generated as an R -module and such that $A \otimes_R R/\mathfrak{m}$ is a central simple R/\mathfrak{m} -algebra for every maximal ideal \mathfrak{m} of R .

Example 6.

- 1) Let B be a central simple k -algebra, and let R be a commutative k -algebra. Then $A = B \otimes_k R$ is an Azumaya R -algebra.

Indeed, since B is finite dimensional over k , $B \otimes_k R$ is finitely generated as an R -module. Let \mathfrak{m} be any maximal ideal of R . Since R is a k -algebra, k identifies to a subring of R , and we have a ring morphism $k \rightarrow R/\mathfrak{m}$ which is injective, since k is a field. Hence R/\mathfrak{m} is a field extension of k . Now we have

$$A \otimes_R R/\mathfrak{m} = (B \otimes_k R) \otimes_R R/\mathfrak{m} \cong_{R/\mathfrak{m}} B \otimes_k R/\mathfrak{m}.$$

Since B is a central simple k -algebra, $B \otimes_k R/\mathfrak{m}$ is a central simple R/\mathfrak{m} -algebra (see [15, Corollary III.1.5 (2)]) and we are done.

- 2) If A and A' are Azumaya R -algebras, then $A \otimes_R A'$ is an Azumaya R -algebra. First, since A and A' are finitely generated as R -modules, so is $A \otimes_R A'$. Now for every maximal ideal \mathfrak{m} of R , we have

$$(A \otimes_R A') \otimes_R R/\mathfrak{m} \cong_{R/\mathfrak{m}} (A \otimes_R R/\mathfrak{m}) \otimes_{R/\mathfrak{m}} (A' \otimes_R R/\mathfrak{m}).$$

This last R/\mathfrak{m} -algebra is the product of two central simple R/\mathfrak{m} -algebras by assumption, hence a central simple R/\mathfrak{m} -algebra by [15, Corollary III.1.5 (1)].

- 3) For all $n \geq 1$, $M_n(R)$ is an Azumaya R -algebra. Indeed, $M_n(R)$ is a finitely generated R -module, and for every maximal ideal \mathfrak{m} of R , we have

$$M_n(R) \otimes_R R/\mathfrak{m} \cong_{R/\mathfrak{m}} M_n(R/\mathfrak{m}),$$

which is central simple over R/\mathfrak{m} .

- 4) We will assume in this example that R is such that for all maximal ideals \mathfrak{m} , R/\mathfrak{m} is of characteristic not 2. Let $a, b \in R^\times$, and consider the R -submodule $(a, b)_R$ of $M_4(R)$ generated by the matrices

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, e = \begin{pmatrix} 0 & a & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & a \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$f = \begin{pmatrix} 0 & 0 & b & 0 \\ 0 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, ef = \begin{pmatrix} 0 & 0 & 0 & -ab \\ 0 & 0 & b & 0 \\ 0 & -a & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Straightforward computations show that these matrices are linearly independent over R , and that we have

$$e^2 = a, f^2 = b, fe = -ef.$$

It easily follows that $(a, b)_R$ is a free R -module of rank 4, which is an R -subalgebra of $M_4(R)$. This R -algebra is denoted by $(a, b)_R$.

It can be viewed also as the R -algebra generated by two elements e, f subject to the relations

$$e^2 = a, f^2 = b, ef = -fe.$$

Then $(a, b)_R$ is an Azumaya R -algebra. Indeed, let \mathfrak{m} be a maximal ideal of R . Since $a, b \in R^\times$, a and b are non-zero elements of R/\mathfrak{m} . The explicit realization above shows easily that we have

$$(a, b)_R \otimes_R R/\mathfrak{m} \cong_{R/\mathfrak{m}} (\bar{a}, \bar{b})_{R/\mathfrak{m}},$$

and it is well known that over a field of characteristic not 2, the quaternion algebra generated by symbols e and f subject to $e^2 = \bar{a}, f^2 = \bar{b}, ef = -fe$ is a central simple algebra. Hence the conclusion.

Azumaya algebras share common properties with central simple algebras. For example, we have the following lemma.

Lemma 20. *Let A and B be two R -algebras. Assume that A is an Azumaya R -algebra, and that B is a faithful R -algebra, that is the R -algebra map*

$$\begin{aligned} R &\longrightarrow B \\ r &\longmapsto r \cdot 1_B \end{aligned}$$

is injective. Then every R -algebra morphism $f : A \rightarrow B$ is injective.

Proof: Let A, B and $f : A \rightarrow B$ as in the statement of the lemma. Then $\ker(f)$ is a two-sided ideal, hence an A - A -bimodule. By [20, Ch. III, Theorem 5.1.1. (2)], A is central, that is the R -algebra map

$$\begin{aligned} R &\longrightarrow Z(A) \\ r &\longmapsto r \cdot 1_A \end{aligned}$$

is an isomorphism, and separable, meaning that A is a projective module for the natural $A \otimes_R A^{op}$ -module structure induced by the multiplication map. By [21, Corollary 3.7], there exists an ideal I of R such that $\ker(f) = I \cdot A$. Since $\ker(f) = I \cdot A$, for all $x \in I$, we have

$$0_B = f(x \cdot 1_A) = x \cdot f(1_A) = x \cdot 1_B.$$

By assumption on B , we get $x = 0$. Thus $I = 0$, and $\ker(f) = 0$. ■

Remark 9. If B is any ring, and R is a commutative subring of B , then the product law endows B with the structure of an R -algebra satisfying the condition of the previous lemma, since for any $r \in R$, we have $r \cdot 1_B = r 1_B = r$.

The following result was proven in [22, Theorem 32], and will be useful to prove the next proposition.

Theorem 11. *Let R be a Hensel ring, with unique maximal ideal \mathfrak{m} . For every central simple R/\mathfrak{m} -algebra B , there exists an Azumaya R -algebra A , unique up to R -isomorphism, such that $A \otimes_R R/\mathfrak{m} \cong_{R/\mathfrak{m}} B$.*

Proposition 21. *Let R be an Artin ring, and A, B be Azumaya R -algebras. Then $A \cong_R B$ if and only if $A \otimes_R R/\mathfrak{m} \cong_{R/\mathfrak{m}} B \otimes_R R/\mathfrak{m}$ for every maximal ideal \mathfrak{m} of R .*

Proof: One implication is trivial. To prove the other one, notice that by Theorem 10, we have a ring isomorphism

$$\varphi : R \xrightarrow{\sim} R_1 \times \cdots \times R_s,$$

for some local Artin rings R_1, \dots, R_s . We then have a 1-1-correspondence between the set of Azumaya R -algebras A and the set of tuples (A_1, \dots, A_s) , where A_i is an Azumaya R_i -algebra, which is given by

$$\begin{aligned} A &\longmapsto (A \otimes_R R_1, \dots, A \otimes_R R_s) \\ (A_1 \times \cdots \times A_s) \otimes_{R_1 \times \cdots \times R_s} R &\longleftarrow (A_1, \dots, A_s). \end{aligned}$$

Moreover, $A \cong_R B$ if and only if $A \otimes_R R_i \cong_{R_i} B \otimes_R R_i$ for $i = 1, \dots, s$.

Let \mathfrak{m}'_i be the maximal ideal of R_i . Then the ideal

$$\mathfrak{m}_i = \varphi^{-1}(R_1 \times \cdots \times R_{i-1} \times \mathfrak{m}'_i \times R_{i+1} \times \cdots \times R_s)$$

is a maximal ideal of R , and the canonical projection $R \rightarrow R_i$ induces a ring isomorphism

$$R/\mathfrak{m}_i \xrightarrow{\sim} R_i/\mathfrak{m}'_i.$$

This yields

$$A \otimes_R R/\mathfrak{m}_i \cong_{R_i/\mathfrak{m}'_i} (A \otimes_R R_i) \otimes_{R_i} R_i/\mathfrak{m}'_i.$$

Hence, by assumption we get

$$(A \otimes_R R_i) \otimes_{R_i} R_i/\mathfrak{m}'_i \cong_{R_i/\mathfrak{m}'_i} (B \otimes_R R_i) \otimes_{R_i} R_i/\mathfrak{m}'_i.$$

Since R_i is a local Artin ring, it is a Hensel ring by Proposition 19. The previous theorem then shows that $A \otimes_R R_i \cong_{R_i} B \otimes_R R_i$. Since this is true for all $i = 1, \dots, s$, we get $A \cong_R B$ as required. ■

As a consequence, we get the following proposition, which will be crucial for our coding considerations.

Proposition 22. *Let k be a number field, and let R be a finite-dimensional commutative k -algebra. For all $a, b, a', b' \in R^\times$, there exist $c, d \in R^\times$ such that*

$$(a, b)_R \otimes_R (a', b')_R \cong_R M_2(R) \otimes_R (c, d)_R.$$

Proof: Notice first that R is an Artin ring by Example 5. Let \mathfrak{m} be a maximal ideal of R . Notice that R/\mathfrak{m} is an extension of k of finite degree, the k -vector space structure being given by the map $k \rightarrow R \rightarrow R/\mathfrak{m}$. Hence R/\mathfrak{m} is a number field (and $(a, b)_R$, etc., are Azumaya algebras over R). Since the exponent and index of central simple algebras over a number field must be equal, and since the exponent of the tensor product of two quaternion algebras over R/\mathfrak{m} is at most 2, the tensor product is of the form $M_2(B)$, where B is either a division algebra of index 2, and hence expressible as a quaternion algebra, or else, B is itself $M_2(R/\mathfrak{m})$, which is expressible as the quaternion $(1, 1)_{R/\mathfrak{m}}$. In either case, therefore, there exists $\bar{c}_\mathfrak{m}, \bar{d}_\mathfrak{m} \in (R/\mathfrak{m})^\times$ such that

$$\begin{aligned} ((a, b)_R \otimes_R (a', b')_R) \otimes_R R/\mathfrak{m} &\cong_{R/\mathfrak{m}} (\bar{a}, \bar{b})_{R/\mathfrak{m}} \otimes_{R/\mathfrak{m}} (\bar{a}', \bar{b}')_{R/\mathfrak{m}} \\ &\cong_{R/\mathfrak{m}} M_2(R/\mathfrak{m}) \otimes_{R/\mathfrak{m}} (\bar{c}_\mathfrak{m}, \bar{d}_\mathfrak{m})_{R/\mathfrak{m}}. \end{aligned}$$

Since R has finitely many maximal ideals by Theorem 10, the Chinese Remainder Theorem shows that there exist $c, d \in R$ such that

$$c \equiv c_\mathfrak{m} \pmod{\mathfrak{m}} \quad \text{and} \quad d \equiv d_\mathfrak{m} \pmod{\mathfrak{m}}$$

for all maximal ideals \mathfrak{m} of R . Notice that $c, d \in R^\times$, since they do not belong to any maximal ideal of R by construction.

For any maximal ideal \mathfrak{m} of R , we then get

$$\begin{aligned} ((a, b)_R \otimes_R (a', b')_R) \otimes_R R/\mathfrak{m} &\cong_{R/\mathfrak{m}} M_2(R/\mathfrak{m}) \otimes_{R/\mathfrak{m}} (\bar{c}, \bar{d})_{R/\mathfrak{m}} \\ &\cong_{R/\mathfrak{m}} (M_2(R) \otimes_R (c, d)_R) \otimes_R R/\mathfrak{m}. \end{aligned}$$

Now apply the previous proposition to conclude. ■

APPENDIX C

CONNECTIONS BETWEEN THEOREM 7 AND THE HURWITZ-RADON-ECKMANN BOUND

In [13], Eckmann provided a solution to the complex version of the Hurwitz-Radon problem (and also described the solution of the original Hurwitz-Radon problem concerning real matrices). Eckmann showed that the maximum number of $n \times n$ complex matrices A_i that satisfy

- 1) $A_i A_j + A_j A_i = 0$ for all $i \neq j$,
- 2) $A_i^2 = -I_n$, and
- 3) $A_i A_i^* = I_n$

is $2t + 1$, where $t = \nu_2(n)$, i.e., the highest power of 2 that divides n . (The original Hurwitz-Radon problem asked for the maximum number of real matrices satisfying these conditions, but with Condition 3 replaced with orthogonality: $A_i A_i^t = I_n$.)

First note that if a matrix satisfies any two of the following three conditions:

$$\begin{aligned} A_i^2 &= -I_n \\ A_i A_i^* &= I_n \\ A_i^* &= -A_i \end{aligned} \tag{7}$$

then it automatically satisfies the third (this is easy to see). If we now compare the hypotheses of Theorem 7 with those of the generalized Hurwitz-Radon problem, we see that Theorem 7 generalizes the Hurwitz-Radon-Eckmann bound in two ways: it does not impose any of the three conditions above in (7) and only considers pairwise anti commutativity, and secondly, it considers the situation where the matrices arise from the embedding of some k -central simple algebra, k a number field, in $M_n(\mathbb{C})$. Since Theorem 7 provides a bound of $2t + 3$, we find that the conditions in (7) drop the possible number by 2.

REFERENCES

- [1] G.R. Jithamithra, B.S Rajan , Minimizing the Complexity of Fast Sphere Decoding of STBCs, *IEEE Transactions on Wireless Communications*, vol 12, no. 12, 2013.
- [2] E. Viterbo, J. Boutros, "A universal lattice decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, 1999.
- [3] E. Biglieri, Y. Hong and E. Viterbo, "On fast-decodable space-time block codes," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, Feb 2009.
- [4] T.P. Ren, Y.L. Guan, C. Yuen, and R.J. Shen, "Fast-group-decodable space-time block code," Proceedings IEEE Workshop (ITW 2010), 2010.
- [5] Chau Yuen, Yong Liang Guan, Tjeng Thieng Tjhung, "On the Search for High-Rate Quasi-Orthogonal SpaceTime Block Code," *Int. J. Wireless Inf. Network*, vol. 13, pp. 329-340, Oct. 2006.
- [6] K. P. Srinath, B. S. Rajan, "Low ML-decoding complexity, large coding gain, full-diversity STBCs for 2×2 and 4×2 MIMO systems," *IEEE J. on Special Topics in Signal Processing: managing complexity in multi-user MIMO systems*, 2010
- [7] N. Markin, F. Oggier, "Iterated space-time code constructions from cyclic algebras," *Information Theory, IEEE Transactions on*, vol.59, no.9, pp.5966–5979, Sept. 2013.
- [8] R. Vehkalahti, C. Hollanti, F. Oggier, "Fast-Decodable Asymmetric Space-Time Codes from Division Algebras," *IEEE Transactions on Information Theory*, vol. 58, no. 4, April 2012.
- [9] L. Luzzi, F. Oggier, "A family of fast-decodable MIMO codes from crossed-product algebras over \mathbb{Q} ," *Proc. IEEE Int. Symp. Inform. Theory*, St Petersburg, July 2011.
- [10] K. P. Srinath, B. S. Rajan, "Generalized Silver Codes," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, Sep 2011.
- [11] L. P. Natarajan, B. S. Rajan, "Asymptotically-Good, Multigroup Decodable Space-Time Block Codes," *IEEE Transactions on Wireless Communications*, vol. 12, no 10, pp. 5035-5047, 2013.
- [12] Lakshmi Prasad Natarajan and B. Sundar Rajan, "Generalized Distributive Law for ML Decoding of Space-Time Block Codes," *IEEE Trans. on Information Theory*, Vol. 59, No. 5, May 2013, pp.2914-2935.
- [13] Beno Eckmann, "Hurwitz-Radon matrices revisited: From effective solution of the Hurwitz matrix equations to Bott periodicity," *Mathematical survey lectures 1943–2004*, Springer-Verlag, Berlin, 2006.
- [14] Grégory Berhuy, Nadya Markin, B.A. Sethuraman, "Fast lattice decodability of space-time block codes," Proceedings of the IEEE International Symposium on Information Theory, Honolulu, Hawaii, 2014.
- [15] G. Berhuy and F. Oggier, An introduction to central simple algebras and their applications to wireless communication, *Mathematical Surveys and Monographs*, Amer. Math. Soc., vol. 191, 2013.
- [16] B.A. Sethuraman, "Division algebras and wireless communications," *Notices of the Amer. Math. Soc.*, vol. 57, pp. 1432–1439, December 2010.
- [17] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high-rate space-time block codes from division algebras," *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2596-2616, Oct. 2003.
- [18] M.F. Atiyah, I.G. Macdonald, *Introduction to commutative algebra*. Addison Wesley (1969)
- [19] J.S. Milne, *Étale cohomology*. Princeton University Press (1980)
- [20] M.-A. Knus, *Quadratic and hermitian forms over rings. Second ed.* Grundle Math. Wiss. **294** (2012)
- [21] F.R. Demeyer, E. Ingraham, *Separable algebras over commutative rings*. Lecture notes in Math. **181**, Springer, Berlin, Heidelberg, New York (1971).
- [22] G. Azumaya, *On maximally central algebras*. Nagoya Math. J., vol. **2** (1951), 119–150.